



中华人民共和国国家标准

GB/T 20092.1—2022

部分代替 GB/T 20092—2013

中文新闻信息内容 第 1 部分：概念模型

Chinese news information content—
Part 1: Concept model

2022-10-12 发布

2022-10-12 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前言	III
引言	IV
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 概念模型	2
4.1 总体结构	2
4.2 文档层模型	4
4.3 稿件层模型	4
4.4 内容项层模型	6
4.5 元数据组描述机制	7
4.6 关系机制	7
4.7 受控词表机制	8

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

本文件是 GB/T 20092《中文新闻信息内容》的第 1 部分。GB/T 20092 已经发布了以下部分：

- 第 1 部分：概念模型；
- 第 2 部分：新闻元数据；
- 第 3 部分：数据交换的 XML 格式。

本文件代替 GB/T 20092—2013《中文新闻信息置标语言》中的第 6 章“CNML 体系结构”，与 GB/T 20092—2013 相比，除结构调整和编辑性改动外，主要技术变化如下：

- 删除了原则（见 2013 年版的 6.1）；
- 将“CNML 体系结构”更改为“概念模型”，并将 2013 年版的有关内容更改后纳入（见第 4 章，2013 年版的 8.1 和 9.1）。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国中文新闻信息标准化技术委员会（SAC/TC 352）提出并归口。

本文件起草单位：新华通讯社、清华大学、国家应用软件产品质量检验检测中心、北京中联合超高清协同技术中心有限公司、北京北大方正电子有限公司、中国科学院计算机网络信息中心、北京智途云天科技有限公司、视觉（中国）文化发展股份有限公司、北京北科光大信息技术股份有限公司、中国人民大学、中国传媒大学、中国大百科全书出版社有限公司、杭州前方信息技术有限公司、东南大学、新华网股份有限公司、中国互联网新闻中心、南方报业传媒集团、重庆日报报业集团、青苹果数据中心有限公司、北京中科大洋科技发展股份有限公司。

本文件主要起草人：武国卫、黄菁、张鹏、王威、王付生、朱建波、朱艳华、丁峰、王骏清、王堃越、许向东、陈锐、张弛、曹学军、刘万福、杨鹏、杨溟、王建平、何崑、张晓伟、张晴、付蓉、路海燕、张贤良。

本文件所代替文件的历次版本发布情况为：

- 2006 年首次发布为 GB/T 20092—2006，2013 年第一次修订；
- 本次为第二次修订。

引 言

GB/T 20092—2013 自发布实施以来,已在很多新闻媒体单位的应用系统中使用。近年来,在标准实施过程中不断涌现一些新的应用场景,产生一些新的应用需求。新闻媒体的部分应用系统也逐渐不再使用 XML 格式,使用 JSON 格式传输稿件的应用越来越多。因为 GB/T 20092—2013 定义的技术内容主要是基于 XML 格式定义的,已经逐渐不适应新闻媒体的应用需求,所以本次修订的主要目的是避免将标准的内容局限于某一种技术,把中文新闻信息内容的概念模式和新闻元数据独立出来,并将新闻信息内容的 XML 格式和 JSON 格式单独进行规范和定义,从而使标准的范围更加明确,定义更加清晰。

GB/T 20092 旨在定义普遍适用于中文新闻信息内容采集、编辑、生成、发布、交换、存储、检索、评估反馈等新闻信息生命周期的各个环节的描述规范,为适应最新技术在新闻媒体中的实际需要,使标准更具有可用性,将标准名称改为《中文新闻信息内容》,拟由四个部分构成。

- 第 1 部分:概念模型。目的在于建立新闻行业中文新闻信息的抽象概念模型,规范中文新闻信息文档的总体结构和描述方式,为新闻内容使用者提供指导。
- 第 2 部分:新闻元数据。目的在于规范描述中文新闻信息内容的各类元数据,便于在新闻媒体行业信息系统中实现信息内容的统一表达,一致描述。
- 第 3 部分:数据交换的 XML 格式。目的在于规定基于 XML 的中文新闻信息数据交换格式,为新闻媒体信息系统的建设提供技术支持。
- 第 4 部分:数据交换的 JSON 格式。目的在于规定基于 JSON 的中文新闻信息数据交换格式,为新闻媒体信息系统的建设提供技术支持。

中文新闻信息内容

第 1 部分：概念模型

1 范围

本文件规定了新闻行业中文新闻信息的抽象概念模型,包括文档层模型、稿件层模型和内容项层模型,以及元数据组描述机制、关系机制和受控词表机制。

本文件适用于新闻信息内容的采集、编辑、生成、发布、交换、存储、检索、评估反馈等新闻信息生命周期的各个环节。也适用于报刊、广播、电视、通讯社、新闻网站、社交媒体等新闻内容提供商及媒体应用与研究机构。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语和定义

下列术语和定义适用于本文件。

3.1

元元数据 meta metadata

定义和描述元数据的数据。

3.2

受控词表 controlled vocabulary

从专业的角度将同一范畴的主题进行归类后形成的词表,并允许以受控的方式进行词条修改和扩展。

3.3

文档 document

符合本概念模型的一篇或者若干篇新闻信息内容数据的集合。

3.4

稿件 item

文档中包含的可管理业务对象,一般有完整的元数据和内容,表示一个完整的新闻信息或其他定义信息。

3.5

内容项 content item

实体稿件类型包含的单一内容,可以包含任何一种文字、图片、图形、音频、视音频、多媒体、应用文档等不同类型的数据内容及其元数据。

3.6

信封 envelop

完成新闻信息交换功能的组件结构。